CUE GENERATION IN CHEMISTRY LEARNING WITH THE BOOST OF TESTING

by

Jiahui Qiu

A Thesis Submitted to the Faculty of the

DEPARTMENT OF EDUCATION

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF ARTS

In the Graduate College

THE UNIVERSITY OF ARIZONA

2019

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis prepared by: Jiahui Qiu
titled:

and recommend that it be accepted as fulfilling the thesis requirement for the Master's Degree.

Jonathan G Tullis — Date: Dec 13, 2019

Adriana D Cimetta — Date: Dec 12, 2019

Eric D Smith — Date: Dec 12, 2019

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

I hereby certify that I have read this thesis prepared under my direction and recommend that it be accepted as fulfilling the Master's requirement.

Jonathan G Tullis — Date: Dec 13, 2019
Thesis committee chair
Educational psychology

Abstract

Cue generation is a common and useful technique for people memorize a variety of materials. Mnemonic cues generated by oneself make difficult things more interpretable and, therefore, boost long-term memory. In order to create effective cues, learners have to predict and generate cues that will match the environmental and mental states at the time of retrieval. Testing is also a powerful technique that support memory for a longer period. The benefits of cue generation and testing effect in memory have been well-established in the literature. Yet few studies focus on the effects of boosting memory of the cues themselves and how that may support learning chemistry information. The goal of the current study is to assess whether and how mnemonic cues help students learn chemistry information. The results of the experiment indicate that generating mnemonic cues benefits chemistry learning more than reading cues that peers have generated. Further, practice retrieving those mnemonic cues during study can improve the recall of the cues, but did not ultimately help students learn chemistry content. Generating mnemonic cues may be a form of deeply encoding the material that effectively boosts student learning in chemistry.

Keywords: Cue generation, testing, metacognition, chemistry learning

Table of Contents

List of Table

List of Figures

Generating one's own cues to support later retrieval has a long history. The first recorded use, a technique called "the method of loci", was found in ancient Geeks to help people remember stories, poems, plays, and lectures (Yates, 2013). By using this technique, people can remember a list by connecting each item to a specific familiar location. In the Middle Ages, the method of loci was revived and occasionally used in mysticism and the occult (Yates, 1966). This method became less popular because of the invention and development of the printing press. From 19^{th} to 20^{th} centuries, however, interest in mnemonics grew again. For example, the first major psychology test relied upon recalling strings of digits using a specific mnemonic device called the "phoneme-digit" mnemonic (James, Burkhardt, Bowers, & Skrupskelis,1890). The "phoneme-digit" mnemonic is a technique utilized to support memorizing numbers by converting numbers into consonant sounds and then into words by adding numbers. In this way, meaningless patterns of digits were transformed into more meaningful words and stories. Mnemonic cues have also been examined in more realistic situations. For example, Atkinson (1975) examined the "keyword" method to teach Russian Vocabulary. With this method, learners can associate the unfamiliar foreign word with a acoustically similar English word to help them memorize. Mnemonic cues were also found to be useful in criminal detection work. Witnesses provided more correct details than those without generating cues or perceiving cues by others in mock-witness tests (Kontogianni, Hope, Taylor, Vrij, & Gabbert, 2018). Students also reported memory cues are beneficial in complex domains that have a large new vocabulary, such as biology (Stagg & Donkin, 2016), physics (Gough, 1977) and chemistry (Banks, 1941). Because chemistry concepts are often abstract, students often have difficulty remembering them. In this study, I examined whether and how mnemonic cues help students learn chemistry information.

Cue generation

Generating cues to remember information is typically effective for improving memory. Generating memory cues may help learners remember information because it relies upon the generation effect. In the generation effect, generating items from memory causes better memory than more passively reading those same items (Slamecka & Graf, 1978). Generation improves memory across a range of tests such as free recall and cued recall, confidence ratings, and cued and uncued recognition (Mäntylä & Nilsson, 1983; Laffan, Metzler-Baddeley, Walker, & Jones, 2010).

Beyond just the generation effect, when learners generate their own cues, they can transform the abstract concept into something that meaningfully connects to their life experiences or existing knowledge. Self-generated cues help learners connect to-be-remember information to their personal experience or long-term memory (Mastropieri, Sweda, & Scruggs, 2000). Also, it encourages learners to transform abstract to-be-remember materials into meaningful chunks (Worthen & Hunt, 2017). Indeed, cues that are generated by oneself support greater retention than cues generated by others (Bloom & Lamkin, 2006). For example, a study focused on timeline interview in crime scenario indicated that participants who used self-generated cues provided more accurate details than those in the other-generated or no cues condition by using mock-witnesses (Kontogianni et al., 2018 More specifically, when learners generate cues for targets, they are more likely to generate distinctive cues that were associated with fewer possible target items than others' cues (Tullis & Benjamin, 2015a). Especially when learners were told their cues would guide them for future retrieval and aware of the related competitors in the to-be-remember materials, the cues they generated were more likely to contain idiosyncratic information and personal experience and associate with fewer potential targets (Tullis & Benjamin, 2015a). For example, when learners were asked to memorize three similar words (e.g., "exam," "quiz," and "test"), they tended to generate cues that were more distinctive and thus reduced the confusions among the target words (Tullis & Benjamin, 2015a). Learners actively and effectively choose and use mnemonic cues base on the study choices because they have the privilege to access their idiosyncratic mental states while outsider such as teachers or peers would not (Jameson, Nelson, Leonisio, & Narens, 1993).

Cue generation and barriers of learning chemistry

The barriers of learning chemistry had been studied for decades. Sirhan (2007) stated that chemistry concepts could be new and conceptually complicated for novice students. The working memory space of the new learners can easily be overloaded which might impede students form storing the chemistry information in long-term memory. Novice learners have difficulty selecting the essential information from less important information (Johnstone & Letton, 1991). By instructing learners to use mnemonic cues, they might connect the chemistry concepts to idiosyncratic information and personal experience and thus decrease the cost of encoding and retrieval, liberating the burden of working memory. Further, with the help of self-generated cues, learners are more likely to generate cues that are more distinctive to distinguish two similar concepts. For example, when students are taught the concepts of "Cations are positively charged ions" and "Anions are negatively charged ions", they are more likely to generate cues to emphasize the difference between "cations" and "anions" instead of focusing on the less important information.

On the contrary, mnemonic cues might be not helpful in learning chemistry. While personal experiences would be helpful in remembering to-be-remember materials, it could be harmful for learning chemistry information. Given the complexity of chemistry, misconceptions have been one of the biggest challenges for students and a large body of studies have been conducted on identification of misconceptions (Gabel, 1999). The cause of misconception comes from learners' previous experience such as information from the world around which are sometimes hard to distinguish from the chemistry information (Bodner, 1991). Learners do not come to learn chemistry with empty minds, and cue generation in this situation may be not useful for them. Additionally, chemistry knowledge is complex and involves abstract concepts. Novices may struggle to effectively connect those abstract concepts to their personal life and to transform those concepts into more meaningful chunks of information.

Creating effective mnemonic cues is also challenging for learners. Successful retrieval depends on the relatedness of the contexts between encoding and retrieval (Roediger & Guynn,1996). In order to create effective cues that support retrieval in the future, learners have to generate cues that will match their environmental and cognitive states at the time of retrieval (Raaijmakers & Shiffrin, 1980). However, the cognitive and environmental contexts can change overtime, so it is important to anticipant one's future circumstance and mental state when generating cues during encoding (Ryskin, Benjamin, & Tullis, 2015). Failure to anticipate one's circumstance and mental state in the future can lead to ineffective cue generation and ultimately failure at retrieval.

Testing effect

Retrieving information supports long lasting memory better than restudying that information; this is known as the retrieval practice effect (Roediger & Karpicke, 2006). Retrieval practice boosts memory for a variety of stimuli, including word lists (e.g., Carpenter & Delosh, 2006), general knowledge facts (e.g., Carpenter, Pashler, Wixted, & Vul, 2008), and foreign language vocabulary (e.g., Carrier & Pashler, 1992). Pyc and Rawson (2010) proposed the mediator effectiveness hypothesis, stating that testing improves memory by supporting the use of more effective mediators during encoding. Therefore, learners might shift their cues to a more effective one during the process of retrieval the cues. A mediator is a word, phrase, or concept that links a cue to a target. Memory for targets increases during testing because mediators generated during testing are more likely to be subsequently retrieved and decoded. Mediator retrieval (mediator is recallable when prompted with the cue) and mediator decoding (mediator elicits the target from memory) are two key factors that make a mediator effective. Given the roles of testing for successful encoding, a secondary aim was to explore whether boosting memory for the cues themselves support retention of information.

The current study

This current study was established to answer three research questions. First, we examined whether generating mnemonic cues helps students learn chemistry concepts. Here, we compared the cue generation in two condition (self-generated cues versus other-generated cues). The reason why we compared these two conditions was that it is closer to real world setting. In the classroom setting, which is common way for students learning chemistry, other-generated cues (e.g., teachers' cues and peers' cues) and self-generated cues are two main sources for students to utilize and remember chemistry facts. College students consistently report generating mnemonic cues to memorize important ideas utilizing rhymes, acronyms, songs, and stories (Van Etten, FreeBern, & Pressley, 1997). Also, students generate mnemonic cues for classes when learning a lot of new facts and terminology (McCabe, Osha, Roche, & Susser, 2013). Previous studies showed that self-generated cues support memory retention than other-generated cues (Bloom & Lamkin, 2006; Tullis & Benjamin, 2015a). We hypothesized that self-generated cues would enhance learning of chemistry facts more than other-generated cues.

Second, we explored whether boosting memory for the cues themselves support retention of chemistry information. In this study, we examined how testing memory for cues affects students' memory for the chemistry content. A large body of studies have shown that retrieval practice boosts memory for variety of stimuli than merely restudy the materials (Carpenter &

Delosh, 2006; Carpenter et al., 2008; Carrier & Pashler, 1992). Further retrieval practice has been shown to boost memory for related information (Carpenter, 2011). So, retrieval practice of the cues themselves may support later memory for the chemistry information. We predicted that learners who received retrieval practice during encoding are more likely to retrieve the mnemonic cues and have higher rate of correctness on recalling the concepts during the test phase.

Third, we explored whether students can accurately predict how well self-generated and other-generated cues affect their learning by measuring learners' judgements of learning (JOLs). JOLs are a measure of students' metacognitive monitoring in which learners predict how well they will remember information. Learners tend to rate tested items as less recallable than restudied items, and they fail to perceive the benefit on memory of testing (Agarwal, Karpicke, Kang, Roediger, & McDermott, 2008; Tullis, Finley, & Benjamin, 2013). Previous studies have shown that learners predict that generating is more difficult and less fluent than reading (Hertzog, Dunlosky, Robinson, & Kidder, 2003). We hypothesized that learners would report higher JOLs when they were in restudy condition other than in the retrieval practice condition. And they would rate higher JOLs when they read other-generated cues than self-generated cues.

Method

Participants

To detect a medium effect (Cohen's d = 0.5), a G*power analysis was conducted using the G*power software with 0.8 power and an alpha of 0.05 (Faul, Erdfelder, Lang, & Buchner, 2007). The G*power indicated that 34 participants should be recruited in this study. We chose a medium effect size because previous study suggests that cue generation has a medium-size effect on test performance (Finely & Benjamin, 2012). Fifty-two participants in introductory

Educational Psychology courses at the University of Arizona participated in day 1 study for partial course credit. Forty-three participants ultimately came back for day 2 study.

Materials

Thirty-two chemistry concepts, and corresponding questions, were derived from the guidelines of the Next Gen Science Standards. The entire list of chemistry concepts and corresponding questions are displayed in Appendix A.

Design

This study utilized a within-subject 2 (self-generated cue vs. other-generated cue) x 2 (retrieval practice vs. restudy) factorial design. Each participant was required to create their own mnemonic cue for a random half of the chemistry concepts; for the remaining half, participants studied cues that prior participants supplied. Further, half of concepts in the own-generated condition were re-studied and half were given a practice test. Similarly, half of the concepts in the other-generated condition were randomly assigned to the restudy condition and half were assigned to the retrieval practice condition.

Procedure

Before the experiment started, the participants read and signed an informed consent form. Participants completed the experiment in a small lab room on a personal computer, while up to 3 other participants also participated. The experiment was programmed in MATLAB with the psychophysics toolbox (Brainard, 1997).

Leaners were instructed to remember a list of chemistry concepts for an upcoming test. They were told they would remember the chemistry facts in two conditions during the learning phase: reading chemistry facts and mnemonic cues generated by others or reading chemistry facts and generating their own cues for the concepts. In the instructions, we provided four different examples of mnemonic cues for a variety of science facts. Instructions specified that

“Your cues can be images that you imagine, phrases you create, rhymes, or other connecting information that helps you remember the concept.” In the first phase, participants studied the 32 chemistry concepts one at a time in a random order in black 40-point Arial font in the middle of the computer screen. The first participant on each computer generated mnemonic cues for every single chemistry concept; their test data are not included because they never saw nor studied prior participants’ cues. All subsequent participants experienced both the self-generated and other-generated condition. For half of the facts during this initial study, learners generated their own cues. For the other half of the concepts, learners read the chemistry fact and the mnemonic cue that a prior learner on the computer had generated. When a concept was assigned to the other-generated condition, the computer displayed the mnemonic cue that was generated by the most recent prior participant.

After generating or reading mnemonic cues for each of chemistry concepts, learners reviewed the chemistry facts by either restudying them with the original cue or by trying the retrieve the cue that was present during the initial study. To retrieve the cues, the participant answered the question “What was the mnemonic cue that was associated with this concept?” In the restudy condition, the chemistry concept was presented to the participant with the cue that appeared during the first phase. Learners made a JOL for each cue that they generated or read when they studied concept during this second phase. For each JOL, they rated how likely they answer the question correctly at the time of the test on a scale of 1 (not at all likely) to 4 (very likely).

After restudying or attempting retrieval for each of the chemistry concepts, participants left the lab. On average, participants returned to the lab 36.65 hours later (*SD* = 15.37). On the second day, participants took a test for the studied information. The list of 32 questions was presented one at a time in a new random order and participants tried to answer each question without receiving the mnemonic cue. Finally, following the test for the chemistry concepts, participants saw the chemistry facts again one at a time and tried to recall the cues that corresponded to each concept. This allowed us to judge whether recall of the cues supported memory, whether learners could decipher their cues, and whether practice retrieval boosted recall of the cues themselves.

Results

Data coding. Student responses to the three open ended questions (retrieval practice answers, final test answers, and memory for the mnemonic) were coded as either correct or incorrect. Two researchers blind to the conditions coded student responses separately and each researcher coded the entire data set on their own. The first author devised a coding rubric for the chemistry questions, and the second author followed the rubric. The rubric is displayed in Appendix B. Initial codes matched for at least 80% of the data for each of the three coded questions (specific agreement rates are shown in Table 1). After coding student responses separately, the two researchers discussed and revised their codes that disagreed until the coding achieved full agreement.

To code the correctness of chemistry test answers, we coded based on whether the participant remembered the all the central components of the concepts. If they answered the correct knowledge about the concept, we coded their response as 1. If they provided no or partial answers to the concept (as shown in the coding rubric), we coded it as 0. For example, when the participants answered the question of "what do combustion reactions produce?", we expected them to provide the key components of "carbon dioxide" and "water". Only both of the components answered correctly can be coded as 1.

When coding whether students correctly recalled their mnemonic during the initial retrieval practice and during the final test, we coded based on whether the participants recalled the cues they learned or generated in the study phase. The mnemonic cues the participants generated were diverse, so we coded them based on three criteria. First, if the mnemonic contained certain specific words (e.g., a name, place, action, etc.), we only counted their mnemonic memory as correct if their answers included those specific words. Second, if the used cue was an event, then the answer should have the same meaning without details missing. Third, the answers should be as complete as the used cues.

Table 1. Coding reliability between two raters in test answers, retrieval practice, and mnemonic recalled.

	Attempt 1
Test Answers	0.80
Retrieval Practice	0.83
Mnemonic Recalled	0.81

Chemistry test performance. Test performance of the chemistry concepts is displayed in figure 1. The major goal of the experiment was to explore whether generating cues and whether testing memory for those cues helped students remember chemistry information. A 2 (cue generation: self vs. other) x 2 (retrieval practice or restudy) mixed-design ANOVA on test performance showed a significant main effect of cue generation condition, $F(1, 38) = 15.12$, $p < .001$, $\eta_p^2 = .29$. Neither the main effect of retrieval practice, $F(1, 38) = .39$, $p = .54$, $\eta_p^2 = .01$, nor the interaction between cue generation and retrieval practice, $F(1, 38) = .55$, $p = .46$, $\eta_p^2 = .01$, reached significance. Results showed that learners remembered more of the chemistry information when they generated their own cues than when they read a peer's cues.

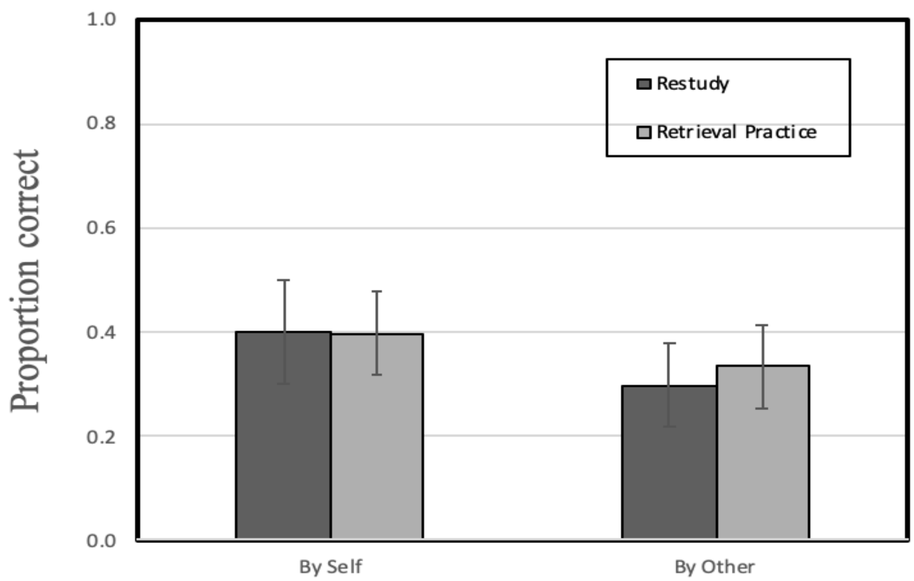

Figure 1. Final chemistry test performance as a function of cue condition and the retrieval practice condition. Error bars indicate 95% confidence intervals for the mean of each column.

An important aspect of effective cue generation is the ability to decode the cues during the test. Decoding a cue means that the learner understands what the cue means and can use the cue to recall the target information. We examined the decodability of the cues across self-generated and other-generated conditions. To do so, we examined the proportion of chemistry questions learners got correct when they were able to recall their cue. When learners remembered their cues, they recalled the same proportion of chemistry facts in the self-generated condition (M = .44, SD = .29) as in the other-generated condition (M = .40, SD = .36), $t(36) = .92$, $p = .36$, Cohen's d = .15. Our data show no significant differences in the decodability of self- and other-generated cues.

Mnemonic recall. Recall of the mnemonic at the end of the experiment is displayed in Figure 2. A two-way repeated measures ANOVA indicated a significant main effect of cue generation condition, $F(1, 38) = 59.88$, $p < .001$, $\eta_p^2 = .61$, a significant main effect on testing, $F(1, 38) = 18.92$, $p < .001$, $\eta_p^2 = .33$, and significant cue generation condition x retrieval practice condition interaction, $F(1, 38) = 12.51$, $p < .001$, $\eta_p^2 = .25$. Follow-up t-tests showed that retrieval practice boosted final recall of self-generated cues ($M = .69$, $SD = .24$) more than restudying ($M = .50$, $SD = .23$), $t(38) = 5.19$, $p < .001$, Cohen's d = .83. However, in the other-generated cue condition, retrieval practice had a smaller impact the final cue recall ($M = .32$, $SD = .21$), $t(38) = .72$, $p = .48$, Cohen's d = .12. The differential impact of retrieval practice on final recall of the cues may depend upon how successfully participants recalled their cues during retrieval practice. Participants successfully retrieved more of their own cues during retrieval practice ($M = .81$, $SD = .19$) than cues of others ($M = .39$, $SD = .22$), $t(38) = 10.74$, $p < .001$, Cohen's d = 1.72.

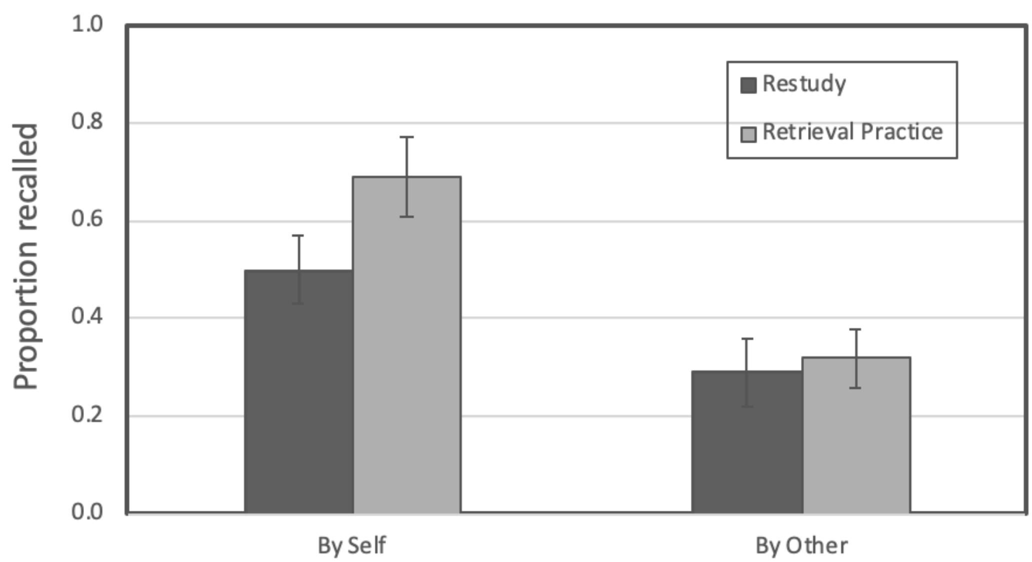

Figure 2. Final cue recalled performance as a function of cue condition and the retrieval practice condition. Error bars indicate 95% confidence intervals for the mean of each column

Judgement of learning. In addition to assessing how much students remembered, we examined how much students thought that they learned. Students' JOLs are displayed in Figure 3. A 2 (cue generation: self vs. other) x 2 (retrieval practice vs. restudy) repeated measures ANOVA on JOLs revealed a significant main effect of cue generation condition, $F(1, 38) = 21.17$, $p < .001$, $\eta_p^2 = .36$, a significant main effect on testing, $F(1, 38) = 5.41$, $p = .026$, η_p^2

= .13, and significant cue generation condition by test interaction, *F*(1, 38) = 13.63, *p* = .001, η_p^2 = .26. Follow-up t-tests revealed that the interaction arose because participants only gave higher JOLs for tested items in the self-generated condition. More specifically, in the self-generated cue condition, learners reported higher JOLs when they were asked to retrieve the mnemonic cues (*M* = 3.28, *SD* = .42) than restudy the cues (*M* = 2.92, *SD* = .50), *t*(38) = 3.88, *p* < .001, Cohen's d = .62. In the other-generated cue condition, there was no statistically significant difference between participants who received retrieval practice (*M* = 2.84, *SD* = .49) and learners who merely restudied the mnemonic cues (*M* = 2.88, *SD* = .51), *t*(38) = -.39, *p* = .70, Cohen's d = .06.

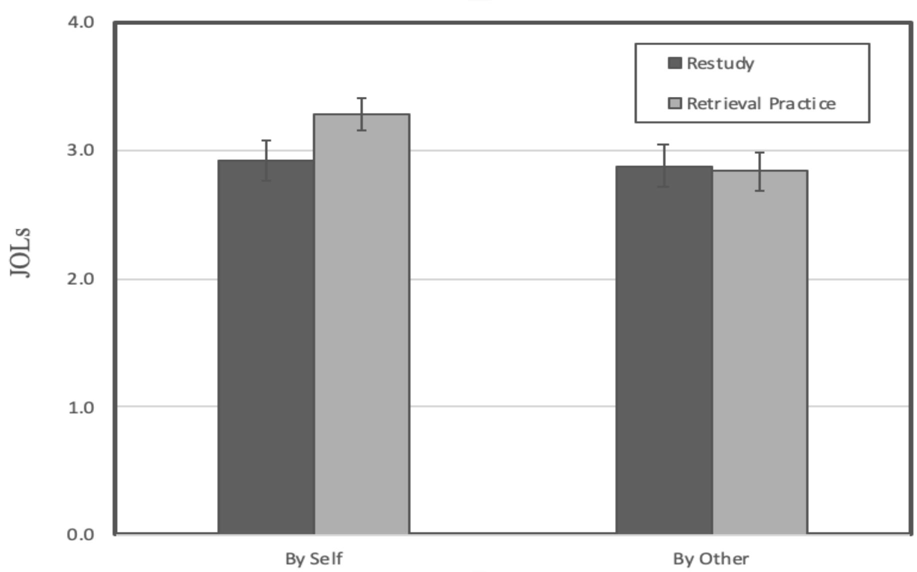

Figure 3. JOLs as a function of cue condition and the retrieval practice condition. Error bars indicate 95% confidence intervals for the mean of each column.

Discussion

In this experiment, I tested whether generating one's own cues supported memory for chemistry concepts and whether retrieval practice interacted with cue generation. First, I tested the performance on the memory of chemistry concepts under the conditions of cue generation and retrieval practice. As predicted, learners who generated chemistry cues remembered more chemistry concepts than learners who received and studied cues that others generated. These data replicate other studies that show self-generated mnemonic cues improve later retrieval across a variety of tasks (Jamieson & Schimpf, 1980; Kuo & Hooper, 2004; Saber & Johnson, 2008; Finley et al., 2012).

Second, I tested the impact of retrieval practice of the memory cues on recall of chemistry content. Practice retrieval of the chemistry cues boosted long-term memory for self-generated mnemonic cues, but this did not translate to better memory for the chemistry content. Retrieval practice of the cue did not improve memory for the cue when learners received and studied someone else's cue. Retrieval practice likely did not bolster long-term retention of other's cues because recall during the practice test was rather low. When trying to retrieve others' cues during a practice test, learners recalled significantly fewer of others' cues. According to the bifurcation model of testing, retrieval practice is only beneficial if you successfully retrieve the target information during the practice test (Kornell, Bjork, & Garcia, 2011). Therefore, if recall is low during the practice test, as it is in the other-generated condition, retrieval practice typically is not as effective as restudying.

Broadly, the data indicate that boosting recall of the mnemonic cues through retrieval practice does not impact memory for the chemistry content. To understand why improving memory for the cues did not bolster memory for the chemistry content, more research is needed. One explanation is that learners may forget to try to use their mnemonic cues during the chemistry test. Learners may not remember that they generated mnemonics for these items and fail to try to retrieve the mnemonics during the chemistry test. Similarly, learners may fail to recognize the value of the mnemonics and not try to retrieve them. Alternatively, when presented with just the test question, learners may not be able to retrieve their cue. In this experiment, I tested final memory for the mnemonic cue by presenting the entire chemistry fact, which may bias my measure of their memory for the cue. Learners may be able to use the fact to help their recall of the cue; when faced with a test of the chemistry content, they may not be able to recall their cue. Future testing is needed to understand the mechanism involved.

Another goal of this study is to examine the learners' metacognition about learning chemistry concepts across the conditions. Learners predicted that self-generated cues would be more beneficial than other-generated cues in helping their later retrieval. According to the encoding fluency hypothesis, learners report greater mnemonic predictions when encoding is easy (Hertzog et al., 2003). In my study, self-generated cues might easier for participants to encode other than reading other-generated cues. Chemistry concepts contain scientific language and the cues generated by others may hard for learners to interpret, while learners can grasp the self-generated cues by their own language. In addition, misconceptions in learning chemistry information are common (Gabel, 1999). Students may have already possessed previous experience, no matter fit or unfit, when they learn the concepts (Bonder, 1991). Therefore, it is likely that they may think others' cues are uninterpretable because they have no such similar experience. For them, self-generated cues are easier to process.

The impact of retrieval practice on learners' JOLs showed unexpected results. Across many prior studies, learners make higher mnemonic predictions for re-read materials than for tested materials. Learners often underrate the benefit of testing and rate restudy as a more memorable way to learn to-be-remembered materials (Agarwal et al., 2008; Tullis et al., 2013). In my data, the impact of retrieval practice on JOLs depended on cue generation condition. When cues were self-generated, learners provided higher JOLs for tested items; when cues were generated by others, no difference existed between JOLs. Mnemonic cues generated often contained idiosyncratic and distinctive experiences, symbols, and connections that may be difficult for other learners to understand. For example, when asking learners to generate cue for the concept of “Electron will fill up orbital shells in this specific sequence: s, p, d, f, g, h, i.” Nearly 90% of the participants made sentence or create a group of words to match the sequence of s, p, d, f, g, h, i. However, these sentences or words are idiosyncratic and more related to personal experience which are not interpretable for other participants. This distinction may become more apparent when learners fail to retrieve the cue during retrieval practice.

Limitations and future directions

As the cues generated by other participants are sometimes hard to interpret and meaningless, I considered to create a list of standard cues that generated by the experts such as school teachers and graduate students. These cues could be easier to understand. The performance on chemistry concepts under the condition of cue generation would be compared. Another limitation of this study, as mentioned above, is the way I tested the memory of cue recall in the final test. It is more appropriate to present the questions to the participants again and ask them the cues they use to answer the questions. Also, it might be necessary to create a circumstance that are closer to the classroom settings, which is that learner can manipulate or control their own learning strategy. Studies show that compared to restrict the learning strategy, allowing learners to control their own learning results in better memory performance (Finley, Tullis, & Benjamin, 2010). In the classroom setting, students can selectively choose to use the self-generated cues or cues that provided by their teachers, and they also can actively choose to restudy or self-test. In this way, we may pinpoint which way of learning are more beneficial to students.

Appendix A

Chemistry Concept	**Question**
A Lewis acid is a compound that can accept an electron pair from a donor compound.	What is a Lewis acid?
A VALENCE electron is located in the outermost shell of an atom and participates in bonds with other atoms.	What are VALENCE electrons?
ALKENES are molecules that only have hydrogen and carbon atoms bonded with at least one double bond.	What are ALKENES?
An Arrhenius acid is a molecule that will donate an H+ ion in solution.	What is an Arrhenius acid?
An ENDOTHERMIC reaction absorbs energy (or heat) from the surrounding system.	What is an ENDOTHERMIC reaction?
An ION is a (positively or negatively) charged particle.	What is an ION?
CATALYSTS lower the activation energy needed to start a reaction.	What do CATALYSTS do?
CATIONS are positively charged ions.	What is a CATION?
COMBUSTION reactions produce carbon dioxide and water.	What do COMBUSTION reactions produce?
Electrons will fill up orbital shells in this specific sequence: s, p, d, f, g, h, i	What is the sequence that electrons fill up orbital shells?
EMULSIFIERS provide a connection between polar and non-polar compounds.	What do EMULSIFIERS do?
FREQUENCY is the number of events in a given unit of time.	What is FREQUENCY?
From longest to shortest wavelength, the electromagnetic spectrum goes: RADIO WAVES, MICROWAVES, INFRARED, VISIBLE, ULTRAVIOLET, X-RAYS, and GAMMA RAYS.	From longest to shortest, what is the order of the waves in the electromagnetic spectrum?
GAMMA RADIATION is a release of a burst of energy (a photon) from an atom.	What is GAMMA RADIATION?
Hydrogen bonds will only form with Nitrogen, Oxygen, and Flourine.	What elements form Hydrogen bonds?
In COVALENT bonds, electrons are shared between two atoms.	What happens in a COVALENT bond?
In the periodic table, chemical elements are organized by increasing number of protons.	How are chemical elements organized in the periodic table?
ISOTOPES have different numbers of neutrons, but same protons and electrons.	What do ISOTOPES differ in?
J. J. THOMSON discovered and identified the first subatomic particle: the electron.	What did THOMSON discover?
OXIDATION reactions involve a loss of electrons.	What happens in OXIDATION reactions?

RUTHERFORD was the first to show that each atom has a nucleus in its center.	What did RUTHERFORD show?
The EMPIRICAL FORMULA for a compound shows the ratio of elements present in a compound (but not the actual numbers of atoms).	What does an EMPIRICAL FORMULA show?
The Pauli Exclusion Principle states that no two electrons can be in the same quantum state.	What is the PAULI EXCLUSION PRINCIPLE?
When acids and bases mix, a SALT is formed.	What forms when an acid and base are mixed?
A BUFFER is a solution that can resist pH change upon the addition of an acidic or basic components.	What kind of solution can resist pH change when an acid or a base is added?
A PRECIPITATE is a solid that forms when two solutions are combined.	What is a solid that forms when two solutions are combined called?
ABSOLUTE ZERO is defined as the point where no more heat can be removed from a system.	What is the point at which no more heat can be removed from a system?
An IONIC bond is formed when one atom accepts or donates one or more of its valence electrons to another atom.	What is IONIC bond?
BOYLE'S LAW describes how the pressure of a gas tends to increase as the volume of the container decreases.	What does the BOYLE'S LAW describe?
NOBLE GAS is stable due to having the maximum number of valence electrons their outer shell can hold.	Why is NOBLE GAS stable?
The IONIZATION energy is the energy required to remove an electron from a gaseous atom or ion completely.	What is IONIZATION energy?
W.F. LIBBY devised a method of estimating the age of organic material based on the decay rate of carbon-14.	What method did W.F. LIBBY devise?

Appendix B

Coding Rubric of Answers

Concept	1	0
A Lewis acid is a compound that can accept an electron pair from a donor compound.	1. Must include the meanings of "electron pair" and "donor". 2. "Electron pair/couple of electrons" is accepted.	
A VALENCE electron is located in the outermost shell of an atom and participates in bonds with other atoms.	1. Must include the meanings of "in the outermost shell" and "bond with other atoms". 2. Other expression can be seen as same as "in the outermost shell": outermost ring. 3. Other expressions that can be seen as same as "bond with other atoms": share, involve in reaction, bond with others.	
ALKENES are molecules that only have hydrogen and carbon atoms bonded with at least one double bond.	1. Must include the meanings of "hydrogen and carbon atoms" and "bond with double bond". 2. Other expressions can be seen as same as "H+ ion": hydrogen ion, hydrogen, H+. 3. Other expressions can be seen as same as "donate": give, take out, loss/lose.	
An ENDOTHERMIC reaction absorbs energy (or heat) from the surrounding system.	1. Must include the meaning of "absorb energy". 2. Other expressions can be seen as same as "absorb": bring in, take, get, input, gain, come into, need.	
An ION is a (positively or negatively) charged particle.	1. Must include the meanings of "charged" and "particle. 2. About "charged": only include "charged" can be counted or both "negatively" and "positively" mentioned. Get or lose electron is ok. 3. Other expressions of "particle": atom, ion.	1. Molecule, element and electron are not correct.
CATALYSTS lower the activation energy needed to start a reaction.	1. Must include the meaning of "lower the activation energy". 2. Other expression can be seen as same as "lower": decrease.	

CATIONS are positively charged ions.	1. Must include the meanings of "positively charged" and "ion". 2. Lose electron(s) is ok.	1. No other expressions can be seen as same as "ion". Atom, element, molecule and particle are not correct.
COMBUSTION reactions produce carbon dioxide and water.	1. Must include the meanings of "carbon dioxide" and "water", nothing else. 2. Other expression can be seen as same as "carbon dioxide": CO2 3. Other expression can be seen as same as "water": H20	
Electrons will fill up orbital shells in this specific sequence: s, p, d, f, g, h, i	1. s, p, d, f, g, h, i in order	
EMULSIFIERS provide a connection between polar and non-polar compounds.	1. Must include the meanings of "connection", "polar and non-polar", and "compound". 2. Other expressions can be seen as same as "connection": mix, combine. 3. Other expression can be seen as same as "compounds": substances.	
FREQUENCY is the number of events in a given unit of time.	1. Must include the meanings of "number of events" and "unit of time".	
From longest to shortest wavelength, the electromagnetic spectrum goes: RADIO WAVES, MICROWAVES, INFRARED, VISIBLE, ULTRAVIOLET, X-RAYS, and GAMMA RAYS.	1. Must in order. 2. Not include "waves".	

GAMMA RADIATION is a release of a burst of energy (a photon) from an atom.	1. Must include the meanings of "release of energy" and "atom". 2. Other expression can be seen as same as "release of energy": explode.	
Hydrogen bonds will only form with Nitrogen, Oxygen, and Flourine.	1. Must include Nitrogen (N), Oxygen (O), and Flourine (F)	1. Somethings else.
In COVALENT bonds, electrons are shared between two atoms.	1. Must include the meanings of "electrons are shared" and "two atoms". 2. Other expression can be seen as same as "electrons": e.	1. Molecule is not correct. 2. "More than two atoms" is not correct.
In the periodic table, chemical elements are organized by increasing number of protons.	1. Must include the meanings of "increasing" and "proton". 2. Other expressions can be seen as same as "proton": atom, atomic.	
ISOTOPES have different numbers of neutrons, but same protons and electrons.	1. Must include "neutrons". Mention of "same protons and electrons" is accepted.	
J. J. THOMSON discovered and identified the first subatomic particle: the electron.	1. Must include "electron".	1. "Valence electron" is not accepted
OXIDATION reactions involve a loss of electrons.	1. Must include the meanings of "loss" and the word "electron" or "e". 2. Other expressions can be seen as same as "loss": reduction, reduce.	

RUTHERFORD was the first to show that each atom has a nucleus in its center.	1. Must include the meanings of "center of atom" and the word "nucleus".	1. "Cell" is not accepted.
The EMPIRICAL FORMULA for a compound shows the ratio of elements present in a compound (but not the actual numbers of atoms).	1. Must include the meanings of "ratio" and "compound". 2. Other meanings can be seen as same as "ratio": most basic form, rate 3. Other meaning can be seen as same as "compound": formula, molecule.	1. "Element" is not accepted.
The Pauli Exclusion Principle states that no two electrons can be in the same quantum state.	1. Must include the meanings of "no two electrons share" and "same quantum state". 2. Other expression can be seen as same as "quantum state": state.	
When acids and bases mix, a SALT is formed.	1. Must include the word "salt".	
A BUFFER is a solution that can resist pH change upon the addition of an acidic or basic components.	1. Must include the meaning of "resist pH change". 2. Other expressions can be seen as same as "resist pH change": same pH, pH equal, balance.	
A PRECIPITATE is a solid that forms when two solutions are combined.	1. Must include the word "precipitate".	
ABSOLUTE ZERO is defined as the point where no more heat can be removed from a system.	1. Must include "absolute zero".	
An IONIC bond is formed when one atom accepts or donates one or more of its valence electrons to another atom.	1. Must include the meanings of "accept or donate electrons" and "one atom to another atom".	

BOYLE'S LAW describes how the pressure of a gas tends to increase as the volume of the container decreases.	1. Must include the meanings of “pressure increase” and “volume of container decrease” or reverse. 2. Other expression can be seen as same as “volume”: size.	
NOBLE GAS is stable due to having the maximum number of valence electrons their outer shell can hold.	1. Must include the meanings of “maximum number” and “valence electrons”. 2. Other expressions can be seen as same as “valence electrons”: outside ring, outer shell. 3. Other expressions can be seen as same as “maximum number”: 8, most possible, all, full.	
The IONIZATION energy is the energy required to remove an electron from a gaseous atom or ion completely.	1. Must include the meanings of “remove an electron” and “gaseous atom or ion”.	1. “Gaseous atom or ion”: “element” is not accepted.
W.F. LIBBY devised a method of estimating the age of organic material based on the decay rate of carbon-14.	1. Must include “carbon-14”. 2. Other expression can be seen as same as “carbon-14”: C-14	

Reference

Agarwal, P. K., Karpicke, J. D., Kang, S. H., Roediger III, H. L., & McDermott, K. B. (2008). Examining the testing effect with open-and closed-book tests. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*(7), 861-876.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American psychologist*, *30*(8), 821.

Banks, B. B. (1941). Chemical mnemonic devices. *J Chem Educ*, *18*, 97.

Bloom, C. M., & Lamkin, D. M. (2006). The Olympian Struggle to Remember the Cranial Nerves: Mnemonics and Student Success. *Teaching of Psychology*.

Bloom, C. M., & Lamkin, D. M. (2006). The Olympian Struggle to Remember the Cranial Nerves: Mnemonics and Student Success. *Teaching of Psychology*.

Bodner, G. M. (1991). I have found you an argument: The conceptual knowledge of beginning chemistry graduate students. *Journal of Chemical Education*, *68*(5), 385.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*(4), 433-436.

Carpenter, S. K., & DeLosh, E. L. (2006). Impoverished cue support enhances subsequent retention: Support for the elaborative retrieval explanation of the testing effect. *Memory & cognition*, *34*(2), 268-276.

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition*, *36*(2), 438-448.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, *20*(6), 633-642.

Finley, J. R., & Benjamin, A. S. (2012). Adaptive and qualitative changes in encoding strategy with experience: Evidence from the test-expectancy paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 632.

Finley, J. R., Tullis, J. G., & Benjamin, A. S. (2010). Metacognitive control of learning and remembering. In *New science of learning* (pp. 109-131). Springer, New York, NY.

Gabel, D. (1999). Improving teaching and learning through chemistry education research: A look to the future. *Journal of Chemical education*, *76*(4), 548.

Gough, W. (1977). The use of mnemonics in mathematics and physics. *Physics Education*, *12*(6), 385.

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 22.

James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology* (Vol. 1, No. 2). London: Macmillan.

Jameson, A., Nelson, T. O., Leonesio, R. J., & Narens, L. (1993). The Feeling of Another Person′ s Knowing. *Journal of Memory and Language*, *32*(3), 320-335.

Jamieson, D. G., & Schimpf, M. G. (1980). Self-generated images are more effective mnemonics. *Journal of Mental Imagery*.

Johnstone, A.H. & Letton, K.M., (1991). Practical Measures for Practical Work, Education in Chemistry, 28(3), 81-83.

Kontogianni, F., Hope, L., Taylor, P. J., Vrij, A., & Gabbert, F. (2018). The benefits of a self-generated cue mnemonic for timeline interviewing. *Journal of Applied Research in Memory and Cognition*, *7*(3), 454-461.

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language*, *65*(2), 85-97.

Kuo, M. L. A., & Hooper, S. (2004). The effects of visual and verbal coding mnemonics on learning Chinese characters in computer-based instruction. *Educational technology research and development*, *52*(3), 23-34.

Laffan, A. J., Metzler-Baddeley, C., Walker, I., & Jones, R. W. (2010). Making errorless learning more active: Self-generation in an error free learning context is superior to standard errorless learning of face–name associations in people with Alzheimer's disease. *Neuropsychological rehabilitation*, *20*(2), 197-211.

Mäntylä, T. I. M. O., & NILSSON, L. G. (1983). Are my cues better than your cues?| Uniqueness and reconstruction as prerequisites for optimal recall of verbal materials. *Scandinavian Journal of Psychology*, *24*(1), 303-312.

Mastropieri, M. A., Sweda, J., & Scruggs, T. E. (2000). Putting mnemonic strategies to work in an inclusive classroom. *Learning Disabilities Research & Practice*, *15*(2), 69-74.

McCabe, J. A., Osha, K. L., Roche, J. A., & Susser, J. A. (2013). Psychology students' knowledge and use of mnemonics. *Teaching of Psychology*, *40*(3), 183-192.

Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, *330*(6002), 335-335.

Raaijmakers, J. G., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In *Psychology of learning and motivation* (Vol. 14, pp. 207-262). Academic Press.

Roediger III, H. L., & Guynn, M. J. (1996). Retrieval processes. In *Memory* (pp. 197-236). Academic Press.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science*, *17*(3), 249-255.

Ryskin, R. A., Benjamin, A. S., Tullis, J., & Brown-Schmidt, S. (2015). Perspective-taking in comprehension, production, and memory: An individual differences approach. *Journal of Experimental Psychology: General*, *144*(5), 898.

Saber, J. L., & Johnson, R. D. (2008). Don't throw out the baby with the bathwater: Verbal repetition, mnemonics, and active learning. *Journal of Marketing Education*, *30*(3), 207-216.

Sirhan, G. (2007). Learning difficulties in chemistry: An overview.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of experimental Psychology: Human learning and Memory*, *4*(6), 592.

Stagg, B. C., & Donkin, M. E. (2016). Mnemonics are an effective tool for adult beginners learning plant identification. *Journal of Biological Education*, *50*(1), 24-40.

Tullis, J. G., & Benjamin, A. S. (2015). Cue generation: How learners flexibly support future retrieval. *Memory & cognition*, *43*(6), 922-938.

Tullis, J. G., Finley, J. R., & Benjamin, A. S. (2013). Metacognition of the testing effect: Guiding learners to predict the benefits of retrieval. *Memory & cognition*, *41*(3), 429-442.

Van Etten, S., Freebern, G., & Pressley, M. (1997). College students' beliefs about exam preparation. *Contemporary Educational Psychology*, *22*(2), 192-212.

Worthen, J. B., & Hunt, R. R. (2017). Mnemonic techniques: Underlying processes and practical applications.

Yates, F. A. (2013). *Art of Memory*. Routledge.